TFE / Master's Thesis Proposal

Modeling Infant Visual Experience: Spatio-Temporal Analysis and Simulation of Visual Input Dynamics Using Deep Learning

Starting date: February/March 2026
Duration: 5–6 months
Location: École Centrale de Lyon, 69130 Ecully, France
Keywords: developmental cognitive science; computer vision; deep learning; SAYCam dataset.
Funding: 650€ / month (net)
PhD follow-up: Yes

Host Laboratories

- **LIRIS – Laboratoire d'InfoRmatique en Image et Systèmes d'Information (CNRS UMR 5205)**
  Team: Imagine
  Supervisors: Zied Bouyahya, Stéphane Derrode
- **ISCMJ – Institut des Sciences Cognitives Marc Jeannerod (CNRS UMR 5229)**
  Team: Cognitive Neuropsychology and Development
  Supervisors: Jean-Rémy Hochmann

Context and Motivation

Early visual experience is fundamental to infant cognitive development. With the advent of egocentric video datasets like SAYCam (Sullivan, J., et al. (2021).), we can now analyze this natural visual environment at an unprecedented scale, from the child's perspective. As an example, the figure below illustrates the data acquisition setup, in which a camera is mounted on the children's foreheads.

This Master's thesis, conducted under the VISIONS project supported by the *Fédération Informatique de Lyon*, aims at developing a computational pipeline towards bridging developmental cognitive science and computer vision. The goal is to leverage state-of-the-art deep learning architectures to characterize the infant's visual world, and model how its spatiotemporal dynamics influence perceptual learning.

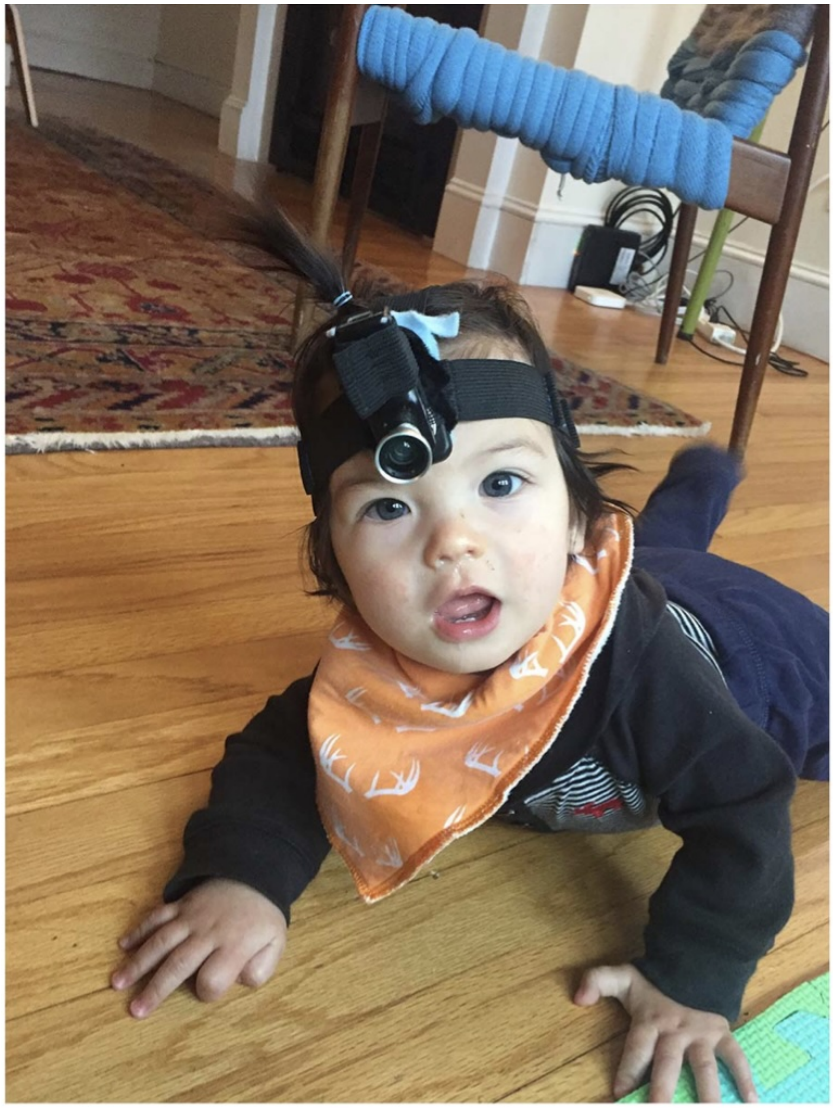

Figure 1. Example of egocentric data acquisition setup (SAYCam dataset).

Objectives

1. **Design an end-to-end video-analysis pipeline** for the SAYCam corpus to detect, classify, and segment salient visual elements (faces, hands, bodies, manipulable objects).
2. **Quantify spatio-temporal dynamics** of the infant's visual input through metrics such as scene-change frequency, motion energy, and object density.
3. **Simulate perceptual regimes** corresponding to early (slow) and later (fast) infant perception, adjusting temporal parameters to reproduce developmental acceleration.
4. **Evaluate and visualize results** to support hypotheses about the relationship between input variability and perceptual learning speed.

Methodology

- **Detection & Segmentation:** YOLO and DETR for object/actor identification; U-Net-based models for pixel-level segmentation.

- **Tracking & Dynamics:** DeepSORT for spatio-temporal consistency; extraction of motion and attention features.
- **Simulation:** Spatio-temporal Transformers and autoencoders to model perceptual encoding under variable frame-rate regimes.
- **Analysis & Validation:** Quantitative comparison between simulated "slow" and "fast" perception and empirical developmental data (Hochmann and Kouider (2022)).

Expected Outcomes

- A **proof-of-concept software pipeline** for large-scale infant-view video analysis.
- Quantitative characterization of the **visual complexity and dynamics** in early life environments.
- Simulation results illustrating how perceptual acceleration may emerge from environmental statistics.
- Eventually, a short **scientific communication or workshop paper** describing the pipeline and findings.

Required Skills

- Python, PyTorch
- Computer vision (OpenCV, YOLO, DETR)
- Data analysis and visualization
- Interest in cognitive modeling and developmental AI

Contact

To apply, please send a CV, your Master grades, and a cover letter to:
zied.bouyahya@ec-lyon.fr **AND** stephane.derrode@ec-lyon.fr.

Applications will be reviewed on a rolling basis, so early contact is encouraged.
Note: Candidates must be referred through the FSD (Fonctionnaire Sécurité Défense) for recruitment.

References

- Bertasius, G., Wang, H., & Torresani, L. (2021). *Is space-time attention all you need for video understanding?* International Conference on Machine Learning.
- Carion, N., Massa, F., et al. (2020). *End-to-end object detection with transformers.* ECCV.
- Hochmann, J.-R., & Kouider, S. (2022). *Acceleration of information processing en route to perceptual awareness in infancy.* Current Biology.
- Sullivan, J., et al. (2021). *SAYCam: A large, longitudinal audiovisual dataset recorded from the infant's perspective.* Open Mind.